Introduction to Data Science

BRIAN D'ALESSANDRO

HEAD OF DS, ZOCDOC

ADJUNCT PROFESSOR, NYU

FALL 2018

COURSE FUNDAMENTALS

INSTRUCTOR

Brian d'Alessandro

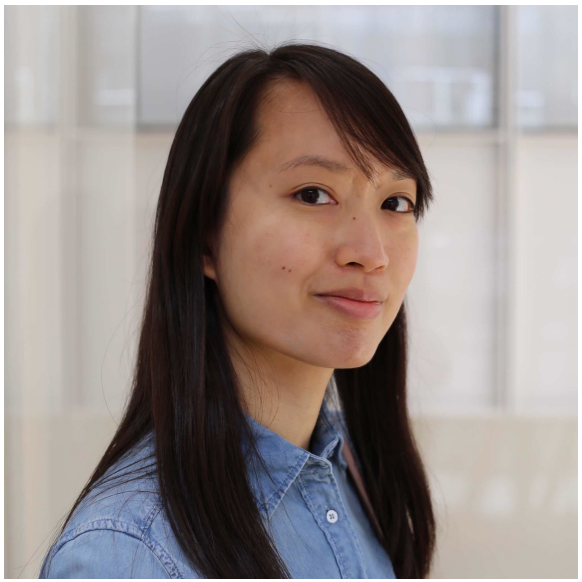
Bio

Education:
Undergrad: Rutgers, Math
Grad: NYU Stern, Statistics

Professional Experience
ZocDoc
Facebook
Dstillery (AdTech)
Meetup.com (Social Web)
American Express (Credit/Risk)

Affiliations/Publications
ACM KDD
Big Data Journal
Machine Learning Journal
SIAM

SECTION LEADERS



Leslie Huang
Email: lesliehuang@nyu.edu

Office hours:
See syllabus

Lee Tanenbaum
Email: leedtan@gmail.com

Office hours:
See syllabus

GOALS OF THIS COURSE

- Understand what a Data Scientist is
- Approach applicable problems data-analytically
- Have hands-on experience mining data

PROGRAMMING!

In order to succeed and participate in this class, You will…

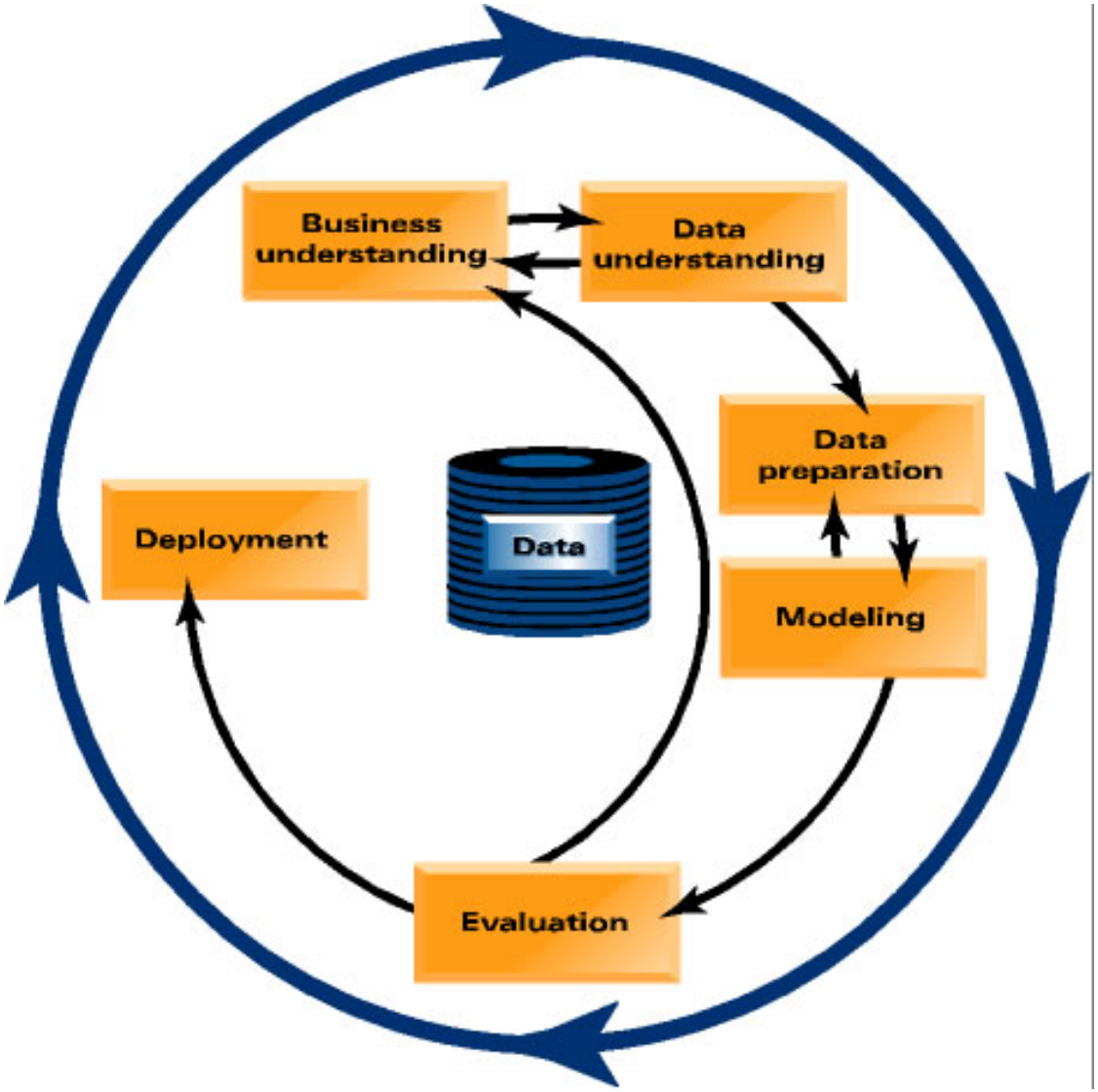
- Need access to a computer with admin privileges
- Have to learn and use the Python programming language.

Please see me after class if this is an issue.

LECTURE OUTLINE

This course will work in the same flow as a typical data mining project.

We'll also peel the layers of data mining like an onion, so the flow might not always be linear.



HOMEWORK

- Will be announced the week they're assigned
  - We'll spend time in class discussing what is expected

- Are all expected to be performed solo
  - Discussions are welcome and encouraged on the course forum

- Will involve Python analysis with discussion of results
  - Open questions will be assigned
  - Code and open answers to be turned in

FINAL PROJECT

The final project will pull together all of the elements you learn from this class and will simulate the experience of being an professional data scientist. Ultimately, we want you to be able to identify a problem, implement a solution, and demonstrate the value of your solution.

Milestones (due dates will be assigned)

1. Choose a team
2. Pick a dataset and a business problem, write a proposal
3. Explore and validate the utility of the data
4. Write a professional report